Detecting Intra-Room Mobility with Signal Strength Descriptors

Konstantinos Kleisouris
CS Dept., Rutgers University
110 Frelinghuysen Rd.
Piscataway, NJ 08854
kkonst@cs.rutgers.edu

Bernhard Firner
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
bfirner@winlab.rutgers.edu

Richard Howard
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
reh@winlab.rutgers.edu

Yanyong Zhang
WINLAB, Rutgers University
671 Route 1 South
North Brunswick, NJ 08902
yyzhang@winlab.rutgers.edu

Richard P. Martin
CS Dept., Rutgers University
110 Frelinghuysen Rd.
Piscataway, NJ 08854
rmartin@cs.rutgers.edu

ABSTRACT

We explore the problem of detecting whether a device has moved within a room. Our approach relies on comparing summaries of received signal strength measurements over time, which we call descriptors. We consider descriptors based on the differences in the mean, standard deviation, and histogram comparison. In close to 1000 mobility events we conducted, our approach delivers perfect recall and near perfect precision for detecting mobility at a granularity of a few seconds. It is robust to the movement of dummy objects near the transmitter as well as people moving within the room. The detection is successful because true mobility causes fast fading, while environmental mobility causes shadow fading, which exhibit considerable difference in signal distributions. The ability to produce good detection accuracy throughout the experiments also demonstrates that our approach can be applied to varying room environments and radio technologies, thus enabling novel security, health care, and inventory control applications.

Categories and Subject Descriptors

C.3 [**Special-Purpose and Application-Based Systems**]: Miscellaneous

General Terms

Experimentation

1. INTRODUCTION

Technology trends have made it possible to add wireless networking into every computing device. Such ubiquity motivates expanding the network for uses beyond communication. For example, a wealth of work over the last ten years has advanced the art of using the network for positioning and localization to the point where such features have been incorporated into commercial products. Other recent investigations have used wireless networks for human presence detection [28], secure key generation [26], and attack detection [4].

In this work, we explore using the wireless network for *mobility detection*. We define mobility as any movement of the transmit antenna. Our definition thus includes changes in transmitter orientation in addition to changes in the spatial position of the device. This paper focuses on *intra-room* mobility. We define this class of events as those of short duration taking place within a room. Such a class of mobility events has a wide range of uses assuming wireless devices are attached to people and objects. A straightforward application is physically securing items. Other uses include health care applications that monitor equipment usage on patients, and retail applications describing items shoppers handle.

Although it is theoretically possible to compute motion by differentiating measured location over time, most current wireless positioning systems are too limited in accuracy to make this approach viable. Indeed, positioning using traditional communication components in realistic scenarios is limited to room scales (10 ft), because of interference and multipath propagation. Hence, it is hard to infer through them whether a device has performed an intra-room movement. Although it is also possible to use a specialized motion sensor, such as an accelerometer, adding sensors increases the device cost, size, and power-consumption, often by the same order of magnitude as the wireless radio. Those reasons preclude the use of accelerometers in tracking applications where sensors must be small or have a very low cost requirements. Also, there are security applications where an accelerometer on a device cannot be trusted since it could be tampered with by an attacker. The wireless radio could be tampered with in the same way, but if the radio is disabled the attack would be detected, and there is no way for an attacker to modify a radio signal so that there is no change in signal strength while a transmitter is actually moving.

In this work, we show we can achieve accurate intra-room mobility detection for different radio technologies under a wide variety of environmental mobility effects. Our ap-

MobiHoc'10, September 20–24, 2010, Chicago, Illinois, USA.

proach relies on the sensitivity of radio channel characteristics to a given spatial antenna position and orientation. We show that our techniques can achieve detection rates of 100% with false positive rates under 2% with little training and in different settings. We show that this accuracy is stable over a wide variety of scenarios, including mobility events performed by different people, at different times of the day, with people walking around the room, and with people moving objects right next to the transmitter. Our experiments are extensive, with close to 1000 mobility events spread out over several months.

Another key result we found is that the parameters needed for our algorithms are very stable. We found a single set of algorithmic parameter constants that can be applied for all scenarios. This is important, because a variety of machine learning approaches require extensive training and may not apply under all scenarios. We found however, that channel stability is easy to discriminate and a single set of thresholds is sufficient. We also show our results apply across different channels and technologies, by using both a Wi-Fi and an active RFID system running in the 900 MHz band [7].

Our technique relies on computing different summaries of Received Signal Strength (RSS) over fixed-time windows, which we call *descriptors*, a term borrowed from the computer vision community. Our technique averages descriptor values collected from a number of receivers or access points (APs) deployed in a room. We explored the performance of three descriptors: standard deviation (σ), absolute difference in the mean (Δ RSS), and histogram distance. We experimented with different histogram distances and concluded that the Earth Mover's Distance (EMD) [21] is the most effective in discriminating changed RSS.

These descriptors have relatively stable values when the environment is static, slightly greater variance when there is movement in the environment, and they exhibit large peaks when the transmitter is actually moving. These observations led us to propose a threshold-based mobility detector. The descriptor threshold value distinguishes between RSS changes when the transmitter is not moving (a stable environment around the transmitter or RSS instability due to perturbations in the radio environment), and changes in RSS due to actual transmitter movement. We also did experiments to test the effects of orientation changes on descriptor values and found them to be similar to the ones during a mobility event, which led us to the conclusion that an orientation change is similar to a transmitter movement. However, in this paper we will focus on mobility detection results for transmitter movement rather than orientation.

The rest of the paper is organized as follows. Section 2 discusses related work. We present our testbed infrastructure, methodology for a series of investigations, as well as our mobility detection algorithm in Section 3. In Section 4, we present our experimental results. Specifically, we show how the different descriptors capture the RSS instability under various conditions, and we give the recall and precision results of the proposed algorithm. Finally, we conclude our work in Section 5.

2. RELATED WORK

There is much work focusing on using specialized sensors and beacons, such as accelerometers, pedometers or motion sensors, to measure different types of human motion [12, 20, 3, 13, 19] and location [12]. Moreover, there has been widespread work [10, 17, 14] on detection and classification of large-scale human mobility patterns based on time-stamped location inferences and a numerical derivative to estimate velocity. [15] proposed a mobility-aware mechanism that adapts the location sampling rate to the target mobility in order to detect when the target crossed a critical region. However, they used an accelerometer to deal with the case in which the target is not moving at constant speeds.

Since specialized hardware is obtrusive and expensive for wide deployments, there is a growing interest in using the existing networking infrastructure to infer mobility patterns. Along this line, [11] classified a user as either moving or stationary based on the variance of a temporally short history of signal strengths from currently the strongest access point. This classification had many transitions between the two states, hence it was smoothened over time with a two-state hidden Markov model (HMM) resulting in an overall accuracy of 87%. [2] monitored GSM signal strength levels and neighboring cell information to distinguish between various states of movement such as walking, travelling in a motor car, and remaining still. The classification of the signal patterns was performed using a neural network and resulted in an average classification accuracy of 80% (driving), 91% (walking), and 96% (stationary). [22] published a similar technique for detecting a user's mobility using signal traces from GSM networks. Their mobility detection system yielded an overall average accuracy of 85% by applying statistical classification and boosting techniques. They extracted a set of seven features to classify the user state as either stationary, walking, or driving. [6] demonstrated the ability to use standard Bluetooth-enabled mobile telephones to recognize social patterns in daily user activity, infer relationships and model organizational rhythms. The information collected included call logs, Bluetooth devices in proximity, and cell tower IDs.

Works closer to ours look at smaller-scale events using wireless LAN or sensor networks and are described in the rest of the section. Our work differs from those below in that we propose a solely RSS-based method, unlike [18] that requires observations at the waveform level, or works that require additional sensors, such as [27].

[18] proposed a location distinction mechanism that used a physical layer characteristic of the radio channel, called temporal link signature, between a transmitter and a receiver to detect when the transmitter or receiver changed position. [27] explored the use of mobile sensors to address the limitation of wireless sensors networks for target detection. [24] proposed a simple motion detection scheme based on RSS readings from a Wireless LAN. They used a variance threshold to declare that a device is moving or stationary. With three access points (APs) and a sliding window of size five samples per AP, they got five percent false positives and about ten percent false negatives. Finally, [16] also presented a motion detection algorithm that was based on the spectral analysis of WLAN radio signal strengths by employing Fast Fourier Transform. A two-state classification scheme was used to deduce if a user is moving or still with an average classification accuracy of 94%.

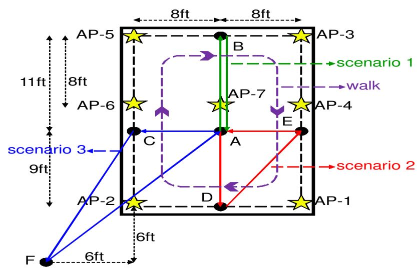
3. METHODOLOGY

In this section we describe our experimental methodology. We first discuss the infrastructure, and then present a series of experiments we conducted to investigate: 1) the impact

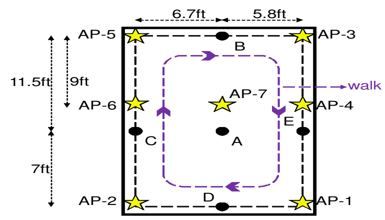
(a) Conference room

(b) Storage room

Figure 1: The two rooms where experiments were carried out. Although the storage room is a much more complex radio environment than the conference room, our techniques worked well in both locations.



(a) Conference room



(b) Storage room

Figure 2: Testbed setups and mobility scenarios. The experiments are described in Section 3.2.

of transmitter orientation on RSS, 2) the impact of local and global environmental instability as well as transmitter movement on RSS, 3) mobility detection of 3 different scenarios. Finally, we present our threshold-based algorithm to detect mobility, and the metrics we used to measure its effectiveness.

3.1 Testbed Infrastructure

We set up two wireless testbeds at our facility and used two types of devices. The first was a Dell laptop equipped with an external NETGEAR MA401 Wi-Fi (IEEE 802.11b/g) card transmitting at a rate of 10 packets/sec. The second was a 900 MHz proprietary active RFID tag (pipsqueak), based on the TI CC1100 digital radio chip [7] with transmission rate of 1 packet/sec. The tag had a simple wire antenna 9.1 cm long, which was set vertical to the floor.

Our experiments were conducted in two rooms that represent two different multipath environments: a conference room (Figure 1(a)), and a storage room (Figure 1(b)). The former was relative empty with a few chairs and a table, while the latter was cluttered with various equipment which thus created numerous multi-path effects. In both rooms we deployed receivers that formed a rectangle measuring 20ft $\times$ 16ft in the conference room, and 18.5ft $\times$ 12.5ft in the storage room, whose locations are depicted by yellow stars in Figures 2(a) and 2(b). Single Wi-Fi and pipsqueak receivers were placed at each location. The black dots in Figures 2(a) and 2(b), i.e. A through F, denote the transmitter locations. Both the receivers and transmitters were placed 3 ft above the floor in the conference room, and 4 ft above the floor in the storage room.

3.2 Experiments

In this section we describe three types of trace-driven [1] experiments we conducted.

Orientation Experiments. The first set of experiments examined the impact of the transmitter orientation on RSS. Specifically, we placed the Wi-Fi transmitter in four different directions in location A of the conference room (Figure 2(a)), and collected signal traces for 3 minutes for each orientation. We then performed the same measurements for the pipsqueak transmitter at the same location but for eight different orientations; four rotations that changed board orientation but not antenna orientation, and four rotations that changed both. The Wi-Fi and RFID transmitters were both rotated by 90° each time.

Stability Experiments. This set of experiments demonstrates how RSS is affected when people move near to and far from the transmitters. This set has two phases, each one consisting of different stages.

In the first phase, the transmitters were placed in location A, and left there for 3 minutes with no one in the room. This stage shows how stable the RSS values are when there is no environmental instability in the room, and we call it the *stability* stage. Then two people walked around the room in a circular fashion near the receivers for 3 minutes. At each moment of the walk, the two people were on opposite sides of the room, and they had normal human walking speed (4-5 ft/sec). This stage captures how RSS is affected by global environmental instability created by people moving within the environment. We call it the *global environmental instability* stage.

Then a person moved a dummy object in close proximity to the transmitters from location A to B. The dummy was 2 feet away from the Wi-Fi and 5 inches away from the pipsqueak in A, and all three formed a line. In this stage only one person was in the room. This stage captures the RSS changes due to local environmental instability, and we call it the *local environmental instability* stage.

Finally, a person moved the two transmitters from location A to B, and placed them next to the dummy object. The transmitters and the dummy had similar distances from each other in B as in A, and they also formed a line. In this stage only one person was in the room. This stage is the *mobility* stage.

Phase 1 continues by repeating the same four stages for locations B, C, D, E. During the third and fourth stages, the dummy and real transmitters were moved from B to C, from C to D, from D to E, from E to A. The duration of phase 1 was approximately 16 minutes. Phase 2 consists of the third and fourth stages of phase 1, for locations A, B, C, D, E, and there was constantly only one person in the room. We collected data by executing phases 1 and 2 in both the conference and the storage room.

In the conference room we collected data in two trials, the first for the RFID tag and the second for both the tag and laptop. On the first trial, the two phases were executed in the order phase 1, phase 2 (4 times in a row), phase 1, whereas on the second trial, they were executed in the order phase 1, phase 2 (3 times in a row), phase 1. The total number of events across both trials included 25 mobility and local instability events, and 10 global instability events for the Wi-Fi transmitter. They also included 55 mobility and local instability events, and 20 global instability events for the RFID tag. In the storage room we collected data for both transmitters in one trial. The two phases were executed in the order phase1, phase 2 (18 times in a row), phase 1. The traces we collected in the storage room included 100 mobility and local instability events, and 10 global instability events for the Wi-Fi transmitter and the RFID tag.

Mobility Experiments. This set of experiments provides a large training set to investigate how well different descriptors detect mobility. All the data for these experiments were collected in the conference room only. We considered the following three mobility scenarios:

1. Scenario 1: The two transmitters start on an edge of the room (B), then move to the center (A), and finally back to the same spot on the edge (B).

2. Scenario 2: The two transmitters start on an edge of the room (E), then move to the center (A), then to a new edge (D), and finally to the previous edge of the room (E).

3. Scenario 3: The two transmitters start on an edge of the room (C), then move out of the room (F), then move into the center of the room (A), and finally back to the same spot on the edge inside the room (C).

We also considered four time periods of stability between movements: 15 secs, 1 min, 3 mins, and 10 mins. Thus, the scenarios include locations in different parts of the room (A-E), as well as outside the room (F), and represent different periods - very short, short, medium and long - of stability at each location. We created data for 100 mobility events that occurred with 15-second and 1-minute intervals, 60 events that occurred with 3-minute intervals, and 50 mobility events that occurred with 10-minute intervals for a total of 310 events per scenario or 930 in total. During the collection of these traces there were no people moving dummy objects, and very little movement in the room except for a single person who was stationary except to move both transmitters from one location to another at the end of each interval. The duration of the movement was 4 to 6 seconds long.

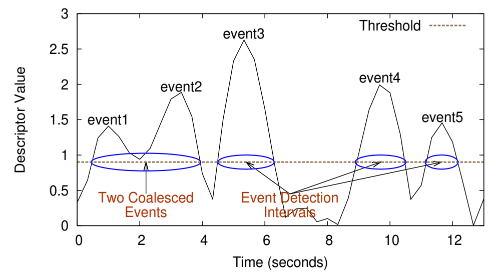


Figure 3: An example of event coalescing with theoretical descriptor values. Identified event start and stop times occur as the descriptor value crosses the threshold.

3.3 Mobility Detection Algorithm

In this section we propose a threshold-based algorithm for mobility detection. The algorithm divides the RSS traces that belong to a given receiver-transmitter pair into windows of fixed-time intervals (e.g. 3, 4, 5 seconds). This is done by using the transmitter's MAC address to extract the RSS values belonging to the transmitter from the traces collected by the receivers. Then it estimates a descriptor using the RSS values of two consecutive time windows. Since the descriptor is estimated based on RSS values that correspond to only one transmitter, our algorithm can be applied for mobility detection of transmitters that move simultaneously, independently from each other.

The algorithm will use the descriptor value and a threshold to decide if a mobility event occured within the two consecutive windows. Since we are using multiple receivers, there are multiple readings for the descriptor value for each pair of consecutive windows that belong to a given transmitter. There are several ways these readings can be compared to the threshold, which leads to several possible voting schemes. In this study, we compare the average of the descriptor values across all receivers to the threshold. Specifically, if the average exceeds the threshold, the algorithm declares there was a mobility event that occured in the midpoint between the two consecutive windows. This voting scheme allows us to apply the same threshold value when using a different number of receivers. All the results in this paper will use the averaging voting method.

To ensure that the detection results are not dependent on the starting time of the first window, the algorithm considers all shifted windows for a shift value s. So, for a window size w and shift size $s = 1$ second, the algorithm considers all windows with starting time 0, 1, 2, $\ldots$, $w - 1$ seconds.

The descriptors we used in this study are: (a) the absolute difference between the RSS mean of the two windows (Δ RSS), (b) the standard deviation of the RSS values across both windows (σ), and (c) the histogram distance between the RSS historgrams corresponding to the two windows. We experimented with different histogram distances, bin-to-bin and cross-bin, and eventually concluded that the Earth Mover's Distance (EMD) [21] (which is cross-bin) is the most effective for the type of analysis we wanted to do (environmental effects on RSS, mobility detection). The EMD is a metric to measure the similarity between two distributions or histograms. Intuitively, given two piles of earth, it gives us how much effort we need to transform one pile into another.

To illustrate the algorithm, we show an example with a window size of 4 seconds and a shift time of 1 second. After 4 seconds, the first time window can be formed but no descriptor values can be estimated. After 8 seconds, two windows can be formed from the intervals (0, 4) and (4, 8), and the first descriptor values can be determined. If the descriptor value exceeds a predetermined threshold, then the corresponding event is marked as starting at *time = 4 seconds*. This means that when using a window of size w, there will always be a latency of $0 \dots w$ seconds in detecting events. If the descriptor value stays above the threshold for several intervals, e.g. (0, 4) (4, 8), (1, 5) (5, 9), and (2, 6) (6, 10), then the interval (4, 6) is a *coalesced interval* and the mobility event is considered to last for three seconds. In this study, we consider shifts of only size $s = 1$ second and, hence, the duration of coalesced events will always be a multiple of seconds. Figure 3 shows a theoretical example of how coalesced intervals are formed.

The metrics we use to evaluate the effectiveness of our algorithm are mobility *recall* and *precision*. Recall gives the percentage of true mobility events that are correctly identified, whereas precision the percentage of predicted events that are correct, and are computed as:

$$Recall = \frac{true\ positive}{true\ positive\ +\ false\ negative}\ , \tag{1}$$

$$Precision = \frac{true\ positive}{true\ positive\ +\ false\ positive}\ . \tag{2}$$

These measures are evaluated on a per event basis and not a per second basis, e.g. a single actual movement event may last for 10 seconds, but our algorithm might only classify it as a movement event after three seconds have already passed - this counts as one true positive, not seven true positives and three false negatives. If a coalesced interval has some overlap with the duration of a true mobility event, then the coalesced interval is considered to be a true positive, otherwise a false positive. In the case of a true positive, we say that the true mobility event was detected by the coalesced interval. If there is no overlap between the duration of a true mobility event with a coalesced interval, then the true mobility event is considered to be a false negative.

There are two different ways to count the number of true positives when calculating recall and precision. The first requires that each mobility event has to be detected separately - two actual mobility events cannot be detected by the same coalesced interval of the algorithm as with $event1$ and $event2$ in Figure 3. This we call *fine-grained* mobility detection. The other viewpoint is that we care whether or not an item has moved, but do not need to know how many times exactly it has moved in a small period of time, and, hence, multiple mobility events can be detected by a single coalesced interval. This we call *coarse-grained* mobility detection. Fine-grained detection is desirable when each individual mobility event is important. For instance, when a chemist must know exactly how many times a dropper has been lifted and used. Coarse-grained detection is desirable when knowing whether or not an item has been moved at all is important. For instance, during a surgery we might care whether an instrument was used, but do not care how many times exactly it was used over a half-hour period. We will see in Section 4 that there is normally little or no difference between the two except in the case where we wish to maximize recall.

In order to evaluate the comparative success of different descriptors, we need to compare one pair of recall and precision values to another. We used the *F-measure* [23] for this purpose, which is defined as:

$$F\text{-}measure = \frac{2 \times precision \times recall}{precision\ +\ recall}\ . \tag{3}$$

The F-measure represents the harmonic mean between precision and recall and is unbiased towards either, thus favoring balanced results (equation 3 does not have any weight to favor precision or recall, and hence precision and recall are treated as equally important). If we were more concerned about either recall or precision, a different measure of comparison would be appropriate.

The thresholds used in our algorithm were determined by a classification algorithm called JRip, a variant of the RIPPER algorithm [5], from the Weka data mining tool [25]. We ran JRip on the local maximums (the peaks) of the data for our three descriptors: standard deviation (σ), absolute difference in the mean (Δ RSS), and Earth Mover's Distance (EMD). The local maximums were used for two reasons. First, the descriptor values during a mobility event will transition from the values seen before the event to the values caused by the event. After the event is over the transition will occur again, but in reverse. The high values observed during the mobility event are really the values of interest to us, and will be captured in the local maximums. The second reason to use the local maximums of the data during training is because it reduces processing time. A period of stability might last for minutes and thus have hundreds of samples, but the threshold just needs to be above the highest points during stability and below the high points during mobility. Training on just the local maximums speeds processing up greatly and does not reduce the quality of the results.

The most successful descriptor was standard deviation. The classification rules of JRip were created using ten-fold cross validation to ensure that they were not over fitted to any particular piece of the data. A ten-fold cross validation splits the input data in a ratio of 9-to-1, where 9 is for training and 1 for validation (we believe that this ratio is typical when validating data sets like ours). No cases of the training were included in the testing. JRip executed 10 runs of cross validations for different 9-to-1 splits, and chose the rules of the run that had the best classification performance.

4. DETECTION RESULTS

In this section we present our results. We first show that our algorithm can cover both Wi-Fi and RFID technologies. We next discuss how the values of descriptors respond to changes in orientation, stability and mobility. Finally, we present recall and precision results for three different types of descriptors.

4.1 The Effects of Window Size and Packet Rate

Figure 4(a) shows the effect of changing the window size for the standard deviation descriptor in the conference room. The window sizes do not alter the shape of the response

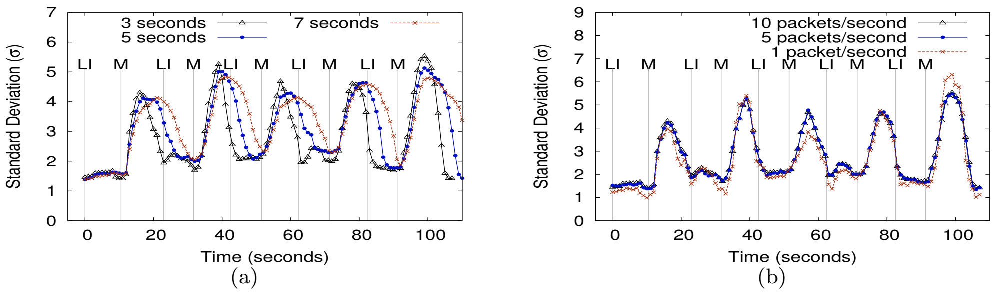


Figure 4: Conference room: The effects of different (a) window sizes with 10 packets per second, and (b) packet rates with a 3-second window from phase 2 of the stability experiments for the Wi-Fi transmitter using standard deviation. The beginning of local instability and mobility events are marked with LI and M respectively.

to mobility, but do alter the width of the curve. Because the rising edge of the curve is the important characteristic, increasing the window size does not improve the detection rates, but rather adds latency and makes it more difficult to distinguish between consecutive short-term events. Results are similar with the other two descriptors. However, longer window sizes smooth noise in the data, and thus do provide some benefit when the number of data samples in a window become small. This suggests that smaller window sizes will give better results for mobility detection, and that the lower bound on the window size will be influenced by the packet rate. We will explore the relationship between window size and event duration more in Section 4.3.

We were also concerned that since the packet rates of the Wi-Fi transmitter (10 packets/sec) and active RFID tag (1 packet/sec) differ, the data might be fundamentally different and the same detection techniques might not be applicable from one dataset to the other. Figure 4(b) shows that this is not the case - the curves are similar under high and low packet rates for the Wi-Fi transmitter. Low packet rates do cause more noise, but this merely suggests that the window size must be large enough to even out the effects of a random process. More generally, the descriptor response being seen is measuring a physical effect, and the packet rate influences how often that response is sampled, but does not change the characteristics of the response. Thus, the packet rate only needs to be high enough so that we observe the width of the physical change.

4.2 Impact of Environmental Changes

Figure 5 depicts the descriptor values over time from phase 1 of the stability experiments in the conference room (see Section 3.2) for both transmitter types. The descriptor values in the figure were determined using 3-second windows, because this single parameter gives good recall and precision for both transmitters. The graphs make it clear why a threshold-based algorithm will be effective for mobility detection. In fact, the curves suggest that three different states are distinguishable using simple thresholds: environmental stability (S), environmental instability caused by motion in a transmitter's global or local environment (GI, LI), and mobility of the transmitter (M).

During the global and local instability states (GI and LI respectively) the descriptor values show greater magnitude and variability when compared to the S state, whereas during the M state they exhibit a peak. Our results also show that local and global instability have similar effects upon signal strength - hence it is hard to distinguish between the two. Moreover, Figure 5(a) shows that, for the Wi-Fi transmitter, standard deviation makes a clearer distinction between states S and LI, GI; the other descriptors show a similar although less pronounced distinction between the states. Although the experiments done for this paper are not an extensive model of environmental instability, our preliminary results suggest it is possible to build a three-state detector.

To show that human motion around a transmitter will not result in a false positive, we refer to Figure 6. It shows that true mobility causes distinct large peaks compared to the peaks caused by the movement of nearby objects. Recall that in the LI stage, a person moves a dummy in close proximity to the real transmitter to another place.

Changes in transmitter orientation were also of interest to us, as we wanted to see their impact on RSS and determine if they could be detected as mobility events. Space prevents us from showing the RSS traces; however, orientation changes are nearly identical to M events. The peaks allow us to conclude that orientation changes are similar to movements of the transmitter. However, in this study, we will apply the mobility detection algorithm we proposed in Section 3.3 for detecting actual M events only.

4.3 Mobility Detection

In this section we evaluate our mobility detection approach. We evaluate the different descriptors on two sets of traces, one for the conference room and one for the storage room. We used the conference room to set the parameters for each descriptor, and then applied those parameters to the storage room traces. Specifically, we created two sets of combined traces: (a) we combined the traces of the signal strength stability experiments and the mobility detection experiments conducted in the conference room, (b) we combined the traces of the signal strength stability experiments conducted in the storage room. We used the data set in (a) above to estimate the precision and recall of our algorithm, and then we applied the same thresholds to estimate the precision and recall of data set (b). Thus, our evaluation allows us to see whether the thresholds estimated in one room can be applied to detect mobility in another.

In the conference room, the stability experiments included 25 actual moves of the Wi-Fi transmitter, and 55 moves of the RFID tag. The mobility experiments, in the same room, included 930 actual moves for both transmitter types. Thus, this data set included over 950 mobility events for testing or training for both the RFID and Wi-Fi transmitters. In the

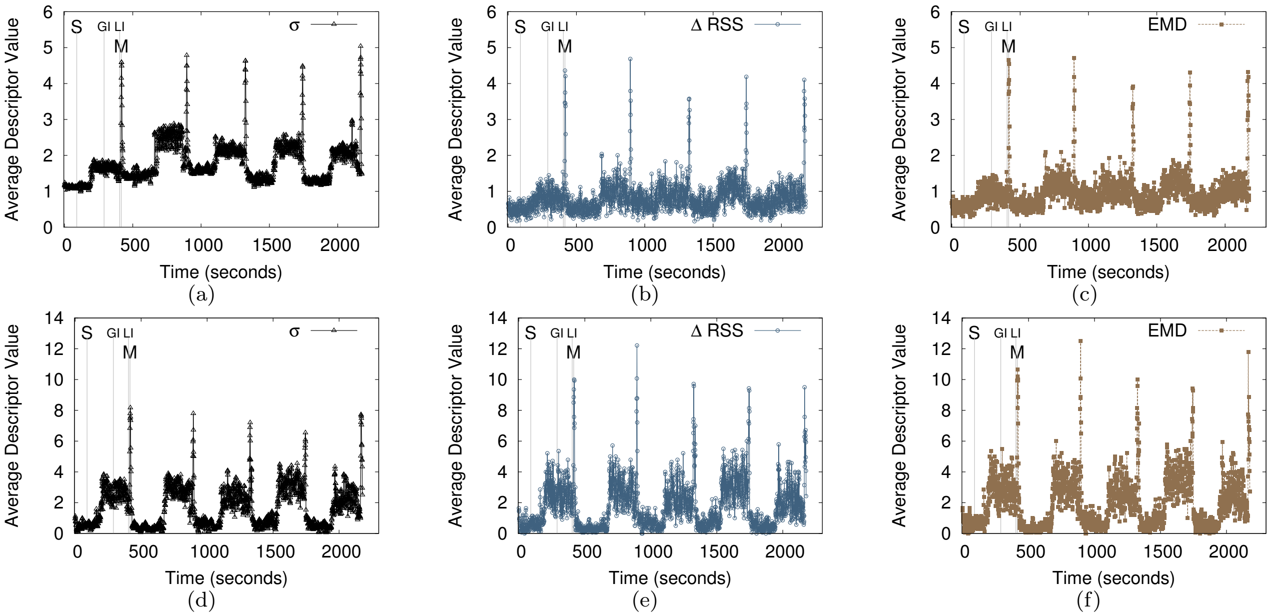


Figure 5: Conference room: Curves for standard deviation (a, d), absolute difference in the mean (b, e), and Earth Mover's Distance (c, f) from phase 1 of the stability experiments with a 3-second window for the (a, b, c) Wi-Fi, and (d, e, f) RFID transmitters. Environmental stability (S), global instability (GI), local instability (LI), and mobility (M) are marked for location A of Figure 2(a).

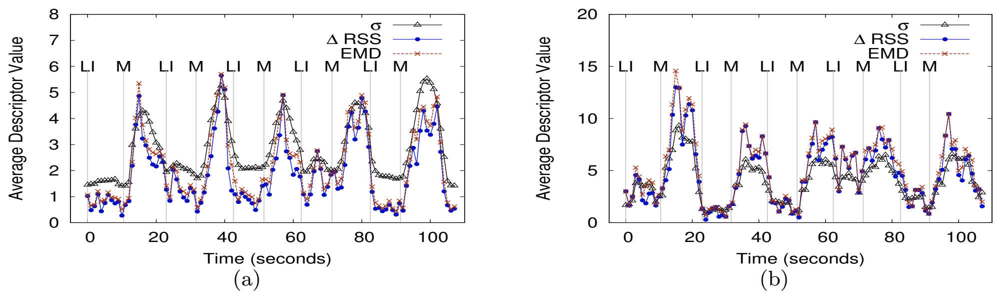


Figure 6: Conference room: Curves for standard deviation, absolute difference in the mean, and Earth Mover's Distance from phase 2 of the stability experiments with a 3-second window for the (a) Wi-Fi, and (b) RFID transmitters. The beginning of local instability and mobility events are marked with LI and M respectively.

storage room, the stability experiments included 100 actual moves for both transmitters. The extensive amount of actual events (close to 1000) and radically different room environments provides confidence that our results are applicable across a wide range of indoor environments.

The optimal F-measure values for mobility detection and corresponding window sizes for the conference and storage room appear in Table 1. These values were estimated for the approximately 1000 actual mobility events, by comparing the F-measures of each descriptor for window sizes from two to ten seconds. The ensemble in this table refers to using all three descriptors at the same time to estimate a threshold. The second entry for the standard deviation descriptor for each transmitter in the conference room shows the results when we maximize recall. Results are for fine-grained precision, which requires that each individual mobility event is detected separately, except for the maximized recall entry for the RFID transmitter in the conference room. Perfect recall requires that our descriptor threshold is set lower than we used for the optimal F-measures, and thus two temporally close mobility events with local instability event in between might cause the descriptor value to stay above the lower threshold through their entire duration and treat M LI M as a single M event. From a fine-grained viewpoint one of the events would be considered missed, so a coarse-grained evaluation is necessary for perfect recall.

The F-measure of the most successful descriptor, standard deviation, with different window sizes is shown in Figure 7 for the conference room. It is clear from the figure that the predictions based upon the descriptor follow the same behavior for both transmitters, although the Wi-Fi and RFID transmitters have different packet rates and operate on different frequencies. This prompts us to suggest that the standard deviation descriptor is applicable across a range of hardware and packet rates. The lower limit on the window size is influenced by the packet rate. The standard

Table 1: The best obtained results for mobility detection as determined by F-measure. An empirically generated result with maximized recall for the standard deviation (second entry) is shown for both RFID and Wi-Fi in the conference room. The results from the storage room serve as a validation set and use the best descriptor (standard deviation) with thresholds determined in the conference room.

Location		Type	Precision	Recall	F-measure	Window Size (seconds)
Conference Room	RFID	σ	0.972	0.986	0.979	3
			0.851	1	0.919	5
		Δ RSS	0.930	0.965	0.947	10
		EMD	0.933	0.981	0.956	5
		ensemble	0.950	0.994	0.977	4
	Wi-Fi	σ	0.984	0.999	0.991	2
			0.980	1	0.990	2
		Δ RSS	0.890	0.970	0.928	7
		EMD	0.915	0.992	0.952	8
		ensemble	0.974	1	0.987	2,3
Storage Room	RFID	σ	.802	1	.890	3
			.789	1	.882	4
			.874	1	.932	5
			.840	1	.913	6
			.846	.991	.912	7
	Wi-Fi	σ	.981	1	.990	3
			1	1	1	4

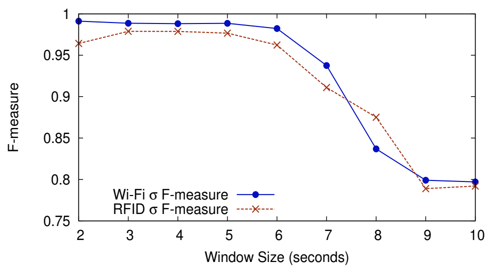


Figure 7: Conference room: F-measure values for the WiFi and RFID transmitters using different window sizes and fine-grained detection.

deviation descriptor is measuring a physical event, so lower data rates will introduce more noise into the result. The Wi-Fi results are better with a smaller window size than the RFID results, because the packet rate is much higher (10 packets/sec vs. 1 packet/sec), and, thus, descriptor values for the Wi-Fi data exhibit less noise.

The F-measure for the Wi-Fi and RFID data begins to fall at six-second windows (Figure 7), because the duration of most of our movement events was between 4 and 6 seconds long. When the window size is larger than the movement event, the window must always contain some stationary RSS values, which gives a smaller separation between mobility and stability events than when the window sizes are small enough to contain only mobility or stability samples.

Our results suggest that a small window size works well when using standard deviation as a descriptor for mobility events for both types of transmitter, which makes sense intuitively - motion will cause variability in the RSS which in turn causes an increase in variance. Not surprisingly, EMD and Δ RSS work better with larger window sizes. This is because these two descriptors compare the signal values and detect mobility by observing the differences in signal strength at different locations; standard deviation uses all the RSS values of two consecutive windows, whereas EMD and Δ RSS build a histogram and estimate a mean for each window separately. There is little to no gain in information in an ensemble for detecting mobility which suggests that, for the same window size, these descriptors carry almost the same information.

Finally, we wished to see how many receivers are necessary to obtain good mobility results. Thus, we computed the F-measure values for every possible combination of receivers using the same descriptor threshold for standard deviation that we determined using seven receivers. The results are plotted in Figure 8(a) for the conference room; each point is a unique combination of receivers. It is clear that more receivers generally lead to better results, but even the worst combination of five receivers for the Wi-Fi data (98.0% F-measure) is within a percent of the performance with all seven (98.9% F-measure). However, the results from the storage room in Figure 8(b) do not show a steady decline in errors. This means that a different factor is at work causing errors in those predictions. Future work needs to be done to determine metrics for the number and placement of receivers in a room. However, with this data, we can see that the probability of error decreases progressively with additional receivers in some environments, and with a reasonable number of receivers, error can be reduced to an acceptable level.

4.4 Location Discrimination

A different goal from mobility detection is *location discrimination*. The idea here is not to detect movement, but rather changes in position. The two goals may or may not be equivalent depending on the latency involved in defining the events. For example, ambiguity can arise when a transmitter is moved and placed back in the same location. In such cases a mobility event has occurred, but a change in location may or may not have occurred depending on the time-granularity used in location discrimination; mobility detection and location discrimination become equivalent as the time-granularity of the two techniques decreases to 0 (real-time).

To better understand the ability of our algorithm to perform location discrimination, we used conference room traces from the mobility experiments with movement intervals of 1, 3, and 10 minutes, along with the phase 1 stability trials in the same room. We ignored the RSS readings during the mobility event, and used two data windows from either side of it in an attempt to compare the stable RSS values of either location. The rest of our data in the conference room was not included, because location discrimination requires

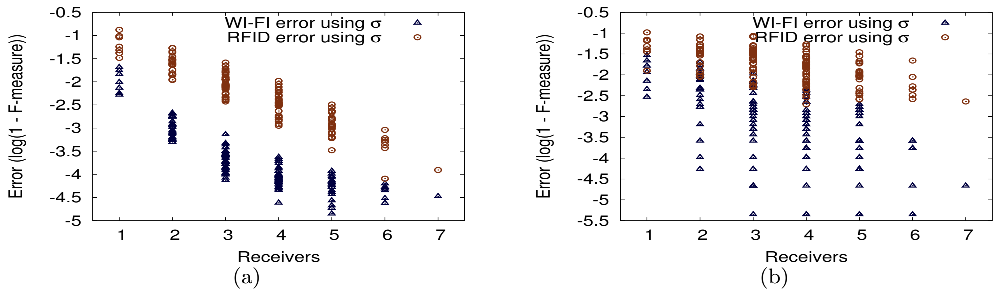


Figure 8: Error for each combination of every subset of the seven receivers for the conference room (a) and the storage room (b).

Table 2: Conference room: The best obtained results for location discrimination using fine-grained detection, as determined by F-measure.

	Type	Precision	Recall	F-Measure	Window Size (seconds)
RFID	σ	0.936	0.782	0.852	8
	Δ RSS	0.949	0.997	0.972	9
	EMD	0.964	0.985	0.974	10
	ensemble	0.942	0.997	0.969	10
Wi-Fi	σ	0.876	0.836	0.855	8
	Δ RSS	0.920	0.984	0.951	10
	EMD	0.922	0.983	0.952	10
	ensemble	0.891	0.994	0.939	9

that the transmitter remains still at the two locations being discriminated.

The results for location discrimination appear in Table 2. The best F-measure for the RFID tag compared to the best from mobility detection (0.974 vs. 0.979) is nearly the same, but now the best result is obtained with EMD. For Wi-Fi the difference between the two results is larger (0.952 vs. 0.991), although EMD again gives the best location discrimination performance. Not surprisingly, standard deviation performs poorly - there is a change in RSS from one window to the next, but the variance of the windows is not large since the transmitters were not moving. Overall, these results suggest that different information is being used in mobility detection and location discrimination.

4.5 Theoretical Foundations

In this section we explore some intuition why we could expect mobility to work drawing on signal propagation theory. Recall that signal propagation can follow two types of distributions: Rayleigh, where the multipath is modeled as the sum of many small components [9], or Rician, where the signal is dominated by a few large components, e.g. when there is a line of sight and a reflection.

We would expect that variance would give good results the more Rician the signal is. Recall that for samples to be uncorrelated they should be more than half a wavelength apart [8], which corresponds to about 16 cm at 902 MHz and 6 cm at 2.4 GHz. Given our M events are larger than these values, we should expect little correlation. In a Rician environment one or more components could change. Because these components are large fraction of the signal, we should see a sharp discontinuity in the response if these are uncorrelated, and hence a large variance. On the other hand, if there are many small components, we would expect them to "average out", and the variance should be less.

However, Rician distributions could increase the false positive rate. For example, if the signal must cross a narrow metal shelf, a person could have a large impact upon the dominant signal when walking through this critical path. Since this sudden fade is similar to a transmitter being moved, such effects are indistinguishable from actual motion. This explains why results are worse in the storage room than in the conference room, as shown in Table 1 and Figure 8.

Future work would measure the actual propagation components to determine their kind and number, and then build signal propagation models appropriate to our environment. Those models could in turn give a stronger theoretic foundation to our results, although such models are beyond the scope of this work.

5. CONCLUSION

Simultaneously tracking indoor objects on a real-time basis is critical for many applications, including security, health care and inventory control. Current approaches that deliver good performance require special hardware, such as motion sensors or cameras, thus making them expensive and intrusive. Low-cost solutions, on the other hand, are very coarse-grained and can only estimate positions spanning large areas and long time periods. In this work, we show how to obtain near perfect seconds-level mobility detection using radios costing a few dollars, which will thus enable new possibilities for these real-time applications.

Our approach is founded on the idea of detecting changes to descriptors, which are summaries of received signal strength measurements over fixed-time windows. We characterized the changes in descriptor values and observed that the changes from transmitter motion differ significantly from the changes from radio and room environmental changes. Our detector is thus based on simple thresholding of the descriptor values. We performed extensive experiments, with close to 1000 mobility events and hours worth of environmental instability, such as people walking in the room and moving dummy objects near the actual transmitter. We found that the standard deviation descriptor with a window size of 2 seconds yielded precision of 98.0% with perfect recall for the Wi-Fi, and the same descriptor with a window size of 3 seconds yielded a precision of 97.2% with a recall of 98.6% for the 900 MHz RFID system. Also, we only needed a modest number of receivers, between 5 and 7, to obtain these results. Finally, we note that we applied this technique in both a cluttered storage room and open conference room, and found very similar results using identical threshold parameters.

6. ACKNOWLEDGMENTS

We would like to thank the anonymous reviewers and our shepherd Dr. Kun Tan for their efforts. This work was partially supported by NSF grants #CNS-0448062 and #CNS-0546072.

7. REFERENCES

[1] Mobility Traces. *http://grail.rutgers.edu/mobilitytraces/*, 2009.

[2] I. Anderson and H. Muller. Context Awareness via GSM Signal Strength Fluctuation. In *The IEEE Pervasive Computing, Late Breaking Results*, pages 27–31, 2006.

[3] L. Bao and S. S. Intille. Activity Recognition from User-Annotated Acceleration Data. In *Pervasive Computing (LNCS)*, volume 3001, pages 1–17, 2004.

[4] Y. Chen, W. Trappe, and R. P. Martin. Attack Detection in Wireless Localization. In *The 26th IEEE International Conference on Computer Communications*, pages 1964–1972, 2007.

[5] W. W. Cohen. Fast Effective Rule Induction. In *The 12th International Conference on Machine Learning*, pages 115–123, 1995.

[6] N. Eagle and A. Pentland. Reality Mining: Sensing Complex Social Systems. *J. Personal and Ubiquitous Computing*, 10:255–268, 2006.

[7] B. Firner, S. Medhekar, Y. Zhang, R. Howard, W. Trappe, P. Wolniansky, and E. Fenson. PIP Tags: Hardware Design and Power Optimization. In *The 5th Workshop on Embedded Networked Sensors*, 2008.

[8] M. J. Gans. A Power-Spectral Theory of Propagation in the Mobile-Radio Environment. *J. Vehicular Technology*, 21:27–38, 1972.

[9] F. Hansen and F. I. Meno. Mobile Fading–Rayleigh and Lognormal Superimposed. *J. Vehicular Technology*, 26:332–335, 1977.

[10] K. Koile, K. Tollmar, D. Demirdjian, H. Shrobe, and T. Darrell. Activity Zones for Context-Aware Computing. In *UbiComp (LNCS)*, volume 2864, pages 90–106, 2003.

[11] J. Krumm and E. Horvitz. LOCADIO: Inferring Motion and Location from Wi-Fi Signal Strengths. In *The 1st Annual International Conference on Mobile and Ubiquitous Systems: Networking and Services*, pages 4–13, 2004.

[12] J. Krumm, L. Williams, and G. Smith. SmartMoveX on a Graph - An Inexpensive Active Badge Tracker. In *UbiComp (LNCS)*, volume 2498, pages 299–307, 2002.

[13] J. Lester, T. Choudhury, N. Kern, G. Borriello, and B. Hannaford. A Hybrid Discriminative/Generative Approach for Modeling Human Activities. In *The International Joint Conference on Artificial Intelligence*, pages 766–772, 2005.

[14] L. Liao, D. Fox, and H. Kautz. Location-Based Activity Recognition using Relational Markov Networks. In *The International Joint Conference on Artificial Intelligence*, pages 773–778, 2005.

[15] T. Lin, P. Huang, H. Chu, and C. You. Energy-Efficient Boundary Detection for RF-Based Localization Systems. *J. Transactions on Mobile Computing*, 8:29–40, 2009.

[16] K. Muthukrishnan, M. Lijding, N. Meratnia, and P. Havinga. Sensing Motion Using Spectral and Spatial Analysis of WLAN RSSI. In *The 2nd European Conference on Smart Sensing and Context*, pages 62–76, 2007.

[17] D. J. Patterson, L. Liao, D. Fox, and H. A. Kautz. Inferring High-Level Behavior from Low-Level Sensors. In *UbiComp (LNCS)*, volume 2864, pages 73–89, 2003.

[18] N. Patwari and S. K. Kasera. Robust Location Distinction using Temporal Link Signatures. In *The 13th ACM International Conference on Mobile Computing Networking*, pages 111–122, 2007.

[19] M. Philipose, K. P. Fishkin, M. Perkowitz, D. J. Patterson, D. Fox, H. Kautz, and D. Hahnel. Inferring Activities from Interactions with Objects. In *The IEEE Pervasive Computing*, volume 3, pages 50–57, 2004.

[20] C. Randell and H. Muller. Context Awareness by Analysing Accelerometer Data. In *The 4th IEEE Computer Society International Symposium on Wearable Computers*, pages 175–176, 2000.

[21] Y. Rubner, C. Tomasi, and L. J. Guibas. The Earth Mover's Distance as a Metric for Image Retrieval. *J. Computer Vision*, 40:99–121, 2000.

[22] T. Sohn, A. Varshavsky, A. LaMarca, M. Y. Chen, T. Choudhury, I. Smith, S. Consolvo, J. Hightower, W. G. Griswold, and E. Lara. Mobility Detection Using Everyday GSM Traces. In *UbiComp (LNCS)*, pages 212–224, 2006.

[23] P. Tan, M. Steinback, and V. Kumar. *Introduction to Data Mining*. Addison Wesley, 2006.

[24] M. Wallbaum and S. Diepolder. A Motion Detection Scheme For Wireless LAN Stations. In *The 3rd International Conference on Mobile Computing and Ubiquitous Networking*, 2006.

[25] I. H. Witten and E. Frank. *Data Mining: Practical Machine Learning Tools and Techniques*. Morgan Kaufmann, San Francisco, 2005.

[26] L. Xiao, J. L. Greenstein, N. B. Mandayam, and W. Trappe. Fingerprints in the Ether: Using the Physical Layer for Wireless Authentication. In *The IEEE International Conference on Communications*, pages 4646–4651, 2007.

[27] G. Xing, J. Wang, K. Shen, Q. Huang, X. Jia, and H. C. So. Mobility-assisted Spatiotemporal Detection in Wireless Sensor Networks. In *The 28th International Conference on Distributed Computing Systems*, 2008.

[28] M. Youssef, M. Mah, and A. Agrawala. Device-Free Passive Localization for Wireless Environments. In *The 13th ACM International Conference on Mobile Computing and Networking*, pages 222–229, 2007.